Accurate Fourier Coefficient and Spectrum Search Algorithm for Short Duration Signals

Mark E. M. Stewart, *Member, IEEE*

***Abstract*—The algorithm presented accurately finds the Fourier coefficients, or spectrum, of discretely sampled, short duration signals. It avoids the FFT's leakage errors. The algorithm uses a classical search approach, namely, bracketing and searching for the frequency, magnitude, and phase of the Fourier modes that minimize signal power—an application of non-linear least squares. The robustness of the procedure is analyzed, showing that bracketing captures only the global minimum and the frequency interval of attraction increases with shorter sampling intervals. The algorithm's operation count and cost are demonstrated. In noiseless test cases, the algorithm can routinely find signal modes to machine precision without leakage. It is validated and compared with the Burgess-Lanczos method. The algorithm is demonstrated by resolving a note from a string instrument and phonemes in human speech. The algorithm has superior accuracy for signals with many modes, and it is valuable for precisely finding the frequency spectrum during a short time interval within a non-stationary signal when the number of signal samples is limited.**

***Index Terms*—Least Squares Methods, Parameter Estimation, Signal Analysis, Speech Processing, Frequency Estimation, Phase Estimation**

Nomenclature

C_f	Fourier coefficient (mode magnitude) at frequency f
f	Frequency of a mode in the signal, Hz
f_s	Uniform sampling frequency of signal, Hz
i	Unit imaginary number
L	Convergence metric, dB
M	Number of cancelling modes
N	Number of DFT samples in sampling interval, $N = T f_s$
$\tilde{N}$	Number of DFT_F samples at test frequency
p	Integer power
P	Signal sample value precision in bits
$r(t)$	Gaussian random noise signal
$s(t)$	Continuous function or signal (single channel)
S_k	Sample value k from discretely sampling signal, $s(t)$,
S_k^M	Sample value k from residual signal, M modes removed
S_{kJ}	Sample value k of polishing signal for mode J
t	Time, s, also $t_k=k/f_s$,
T	Integration time or sampling interval duration, $T = N/f_s$, s
$\tilde{T}$	Integration time at test frequency for DFT_F, s
α	Angular frequency, $\alpha = 2 \pi f$, radians/s
Δt	Time step between samples, $\Delta t = 1/f_s$, s
ε, δ,	Perturbation parameters for frequency, magnitude
ϕ	Perturbation parameter for phase
Φ	Function to be minimized
θ_f	Phase angle of mode that ensures C_f is real, radians
$\lfloor x \rfloor$	Floor function, x
j	Integer index of cancelling mode, from set of M,
J	Integer index of cancelling mode polishing focus
k	Integer index of discrete time, $k\Delta t$; also a general index
ℓ	Integer index of FFT mode, $f_\ell = \ell f_s / N$
^	Accent indicates complex value, otherwise real

Manuscript received May 4, 2013; revised.

Mark Stewart is with NASA Glenn Research Center, Cleveland, OH 44135 USA. (e-mail: Mark.E.Stewart@nasa.gov).

I. Introduction

Finding the frequency and magnitude of a signal's modes is a common problem in science and engineering. Currently, the Fast Fourier Transform (FFT) is a commonly used method. The FFT accurately represents the discrete signal and is very robust and efficient. Yet, in general, it does not provide precise spectrum information, including Fourier coefficients of the underlying signal modes.

The FFT suffers from 'leakage'—additional modes in the spectrum beyond the underlying signal itself. For long duration, stationary signals, a sufficient number of signal samples exist so that FFT results can be averaged to better resolve frequencies. However, non-stationary signals do not have the time duration and number of samples, N, to achieve frequency resolution with an FFT.

The current paper demonstrates a robust algorithm for accurately finding an auditory signal's Fourier coefficients and spectrum. Leakage errors are avoided, and the method can be used in a short interval when an auditory signal is nearly stationary and the number of signal samples is limited. The algorithm is compared with the Burgess-Lanczos method [1].

The algorithm can be valuable beyond auditory signals, in particular, when the precise frequency, magnitude, or phase of signal modes is required.

Section I sets the background and notation for non-specialists. Section II outlines the algorithm. Section III is a perturbation analysis of bracketing. Section IV concerns important properties, and Section V provides examples.

A. Background

In an early example of spectrum analysis from the 1860's, Helmholtz [2] used Helmholtz resonators of glass and later metal to detect individual tones in musical instruments, speech, and other vibrating objects. A Helmholtz resonator

uses sympathetic resonance to detect vibration frequencies. Helmholtz explained the fundamental tone and 'harmonic upper partial tones' by Fourier series.

Fourier [3] used his namesake series to solve the differential equation for heat flow in objects. Fourier's contribution was to recognize that a series of sine/cosine modes can represent almost all functions and signals. Further, these series form solutions to important differential equations, including the wave and heat equations.

Importantly, the wave equation models the underlying physics of many sound sources, plus sound transmission. A string, drum, or beam's vibration are classic examples; sound propagation is another. Linearity of the equation and constant coefficients allow superposition of modes. Hence, the solutions to the wave equation are a sum of sine/cosine modes, and Fourier series represent the vibrations of many sound sources and their propagation. Yet, for those cases (fricative, plosive phonemes) where the physics of the sound source does not result in a wave equation model, Fourier series can still represent the signal.

Fourier series were first calculated manually, and a need has existed for accurate machines and computationally efficient algorithms that automate calculation. Michelson and Stratton [4] devised a mechanical harmonic analyzer and synthesizer that improved the accuracy of Lord Kelvin's machine for analyzing tides [5].

In the earliest years of digital computers, when computers were at least three orders of magnitude slower than today, the Fast Fourier Transform (FFT) was devised to give the Fourier coefficients of a discretely sampled signal. While Cooley and Tukey [6] brought the FFT into wide use in software and hardware, others had independently developed and used similar techniques. Cooley [7] provides a history. Soon after Cooley and Tukey's paper, FFT based spectrum analyzers replaced [8] frequency shifting a signal into a band-pass filter. Contemporary research includes improving discrete Fourier transform (DFT) performance using a reduced distribution of signal samples [9].

One problem in finding a signal's spectrum is that DFTs usually introduce 'leakage'—additional modes in the spectrum beyond the underlying signal modes. Fig. 1 shows a signal with three modes and its FFT. The mode at f = 172.2656 Hz corresponds to an FFT frequency, $4\ f_s/N$, and is found by the FFT without leakage. The mode magnitude is found exactly and the neighboring Fourier coefficients are at the background spectrum. The two other modes are not at FFT frequencies, and they show a local peak, but they also show leakage or non-zero Fourier coefficients across the spectrum. The current algorithm avoids the problem of leakage, since it can search at frequencies between the fixed FFT frequencies.

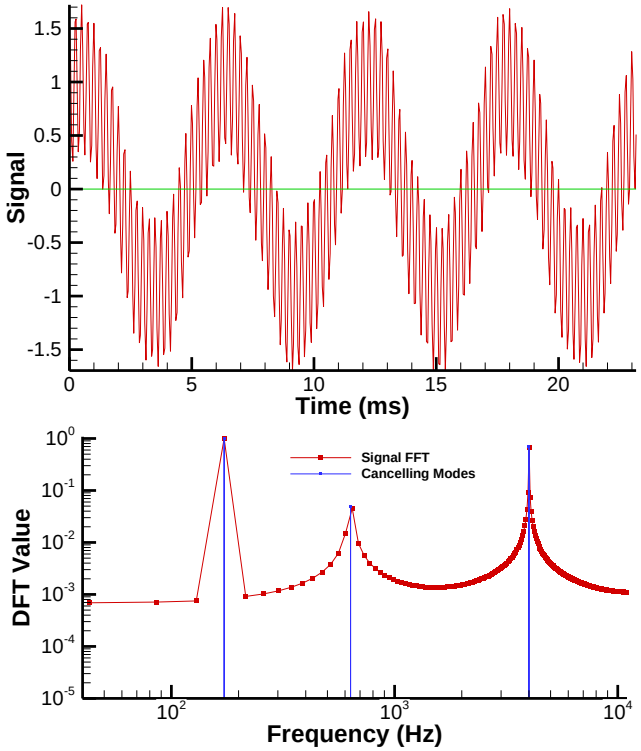

Fig. 1. Noiseless signal (upper) and its FFT (lower) consists of three modes f = 172.2656 Hz, 635 Hz, and 4000 Hz; C_f = 1.0, 0.05, 0.7; $\theta_f = 0.3\pi, 0.2\pi, 0.5\pi$. The mode at f = 172.2656 Hz corresponds to an FFT mode frequency and is found by the FFT *without* leakage. The two other modes, not at FFT mode frequencies, show leakage or non-zero Fourier coefficients spread across the spectrum (lower). Green line above is residual signal, S_k^M; blue histogram lines below are the Fourier modes found by the current algorithm. Throughout the paper, unless stated otherwise, signals are sampled at f_s = 22.05 kHz, N = 512, T = 23.22 ms, P=32 bit resolution; signals are assumed to be band limited and without aliasing.

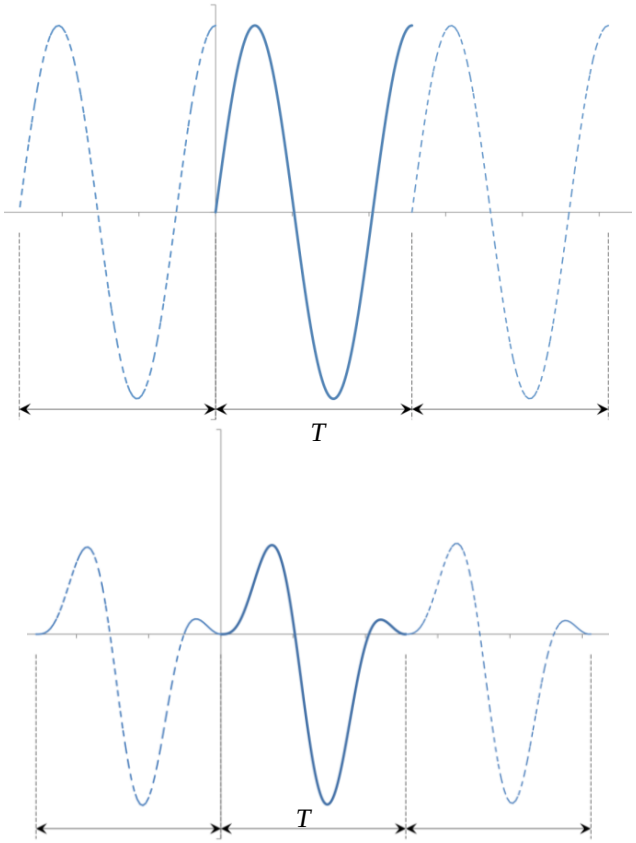

Fig. 2. Periodic extension of a function (upper) and after Hanning filtering (lower). The function is periodic, but not of period T. A discontinuity is created above, and the filter removes it by distorting the signal. A slope discontinuity can also be created in a periodic extension.

Leakage is not an error; the FFT does exactly what it should do. FFT modes are all periodic with period *T*; consequently the series must also be periodic with period *T*. However, if the sampled signal is not periodic with this period, a discontinuity is introduced (Fig 2). In particular, when the signal is periodically extended beyond the interval and has a discontinuity or slope discontinuity, the FFT places a discontinuity where intervals meet—a departure from the intended signal [1]. Data windowing functions reduce FFT leakage by reducing this discontinuity; the sample values are zeroed toward the interval ends, but the signal is distorted.

A statistical estimation approach to reducing the influence of leakage and noise, thereby improving an FFT spectrum, is to find an average value: divide a stationary, long duration signal into sub-intervals, apply an FFT to each sub-interval, and average the magnitude squared Fourier coefficients, $\left|\hat{C}_{f_\ell}\right|^2$, at each frequency; in this manner, the variance of signal's power spectrum (periodogram) is reduced [10] [11]. This procedure is often attributed to Bartlett [12]. For example, Lewicki [13] averaged fifty 1024-point FFTs to resolve the vibration spectrum of a helicopter transmission. For short duration signals, insufficient data exists for this approach.

The Burgess-Lanczos method [1] [14] takes a different approach to circumvent leakage. Instead of minimizing leakage, the method assumes an underlying mode exists and approximates it from adjacent FFT mode parameters.

Other methods exist, beyond the FFT. Instead of a Fourier transform, non-linear least squares (NLLS) and a minimization search can be used to find signal parameters. Friedlander and Francos [15] present theoretical analyses of using NLLS to estimate signal parameters. Angeby [16] applies NLLS and the Gauss-Newton method to polynomial phase signals (PPS), and notes that searching on small, sliding intervals eases numerical problems. For signal parameter estimation in the presence of noise, NLLS is considered statistically efficient. Further, it can be applied to non-uniformly sampled signals.

The all-poles method (maximum entropy method (MEM), autoregressive model (AR), linear predictive coding (LPC)) is another approach to finding spectral content. By using a rational function (1) with poles, sharp spectral features can occur at real frequencies and represent discrete modes in a signal. The method can produce remarkable results, but it is considered sensitive to the model's order and captures noise. In speech recognition [17] [18], the MEM is a popular method of representing a short duration signal due to its speed and simplicity.

$$\frac{a_0}{1 + a_1 z^{-1} + a_2 z^{-2} + \ldots + a_p z^{-p}} \quad . \tag{1}$$

Since the 1980's, wavelets [19] have been used to represent functions as a series, with a transform to find the series—not unlike Fourier series. Wavelets are appropriate for non-stationary and non-periodic signals, as well as functions with discontinuities and sharp peaks. A fast, recursive algorithm [20] exists for finding the discrete wavelet transform.

In electrical hardware, the Phase-Locked Loop (PLL) finds a frequency in a signal. In demodulating an FM radio signal, the frequency is found by a PLL or the Foster-Seeley discriminator.

The human cochlea is a spectrum analyzer for auditory signals; Helmholtz [2] recognized this. Von Békésy [21] demonstrated that sound frequencies resonate at different points along the cochlear partition since decreasing width changes its resonant frequency. Measurements indicate individual frequencies are not sharply resolved in the feline cochlea [22], although for some echo locating bats with long duration calls, resolution may be an order of magnitude sharper.

B. Fourier Series, Coefficients, and Transform

The Fourier series and transform are outlined here for completeness and clarity of notation. The Fourier series is a representation of a continuous function or signal, $\hat{s}(t)$. The Fourier series is expressed in terms of Fourier coefficients, $\hat{C}_{f_\ell}$, and modes, $e^{2\pi i f_\ell t} = \cos(2\pi f_\ell t) + i \sin(2\pi f_\ell t)$, on *[0,T]* as,

$$\hat{s}(t) \approx \sum_\ell \hat{C}_{f_\ell} e^{2\pi i f_\ell t} \quad . \tag{2}$$

Modes come in pairs, $\pm f_\ell$. Here we focus on real signals, *s(t)*, and the real part of the series, namely

$$s(t) \approx \sum_\ell C_{f_\ell} \cos(2\pi f_\ell t + \theta_\ell), \quad C_{f_\ell} = 2\left|\hat{C}_{f_\ell}\right|, \quad f_\ell \geq 0 \,. \tag{2b}$$

A constant, real phase, θ_ℓ, is introduced to make magnitude, C_{f_ℓ}, real

By the Fourier Expansion Theorem [23], the series in (2) will converge to *s(t)* in an interval, if *s(t)* is piecewise continuous in the interval with a finite number of corners, that is, slope discontinuities. *s(t)* must also be periodic outside the interval (a sufficient condition, but not necessary). The coefficients $\hat{C}_{f_\ell}$ are given by (3) in the interval [*0, T*].

The wave equation describes the vibration of stringed instruments and drums and is solved mathematically with a Fourier series. Consequently, a few harmonic modes describe the sound. Similarly, the vibrating human vocal folds generate harmonic modes. Many objects and machines have natural frequencies—due to their shape, restraint, and elastic properties—which define their resonance and vibration in terms of Fourier modes.

In contrast, the fricative phonemes of human speech (voiceless 's' and 'f') are created by turbulent airflow between the tongue and teeth; there are multiple flowpath locations where flow separation can occur. Turbulent airflow involves transient eddies of varying sizes that produce sounds at random frequencies within a frequency range. The mathematical analysis of this sound source does not yield Fourier series solutions, but with enough modes and the Fourier Expansion Theorem, sound representation is possible. Similarly, Fourier series represent the plosive phonemes of human speech ('b', 'p', 't') and their abruptly starting air flow

and pressure wave. Section V gives examples.

The Fourier Transform is a way of finding Fourier coefficients and the series. Under appropriate conditions, the Fourier coefficients, $\hat{C}_{f_\ell}$, of a signal, *s(t)*, are given by,

$$\hat{C}_{f_\ell} = \frac{1}{T}\int_0^T e^{2\pi i f_\ell t}\, s(t)\, dt \qquad (3)$$

where *T* is an integration time that must be an integer number of periods at the frequency, f_ℓ. Further, *s(t)* must be periodic in the interval [*0,T*] or zero outside of it—otherwise leakage can result.

The FFT is a practical method of efficiently finding Fourier coefficients using a discrete approximation to the Fourier Transform (3).

$$\hat{C}_{f_\ell} \approx \frac{1}{N}\sum_{k=0}^{N-1} e^{2\pi i f_\ell t_k}\, s(t_k) \qquad (4)$$

where $t_k = \frac{k}{f_s}, \ell = 0, \pm\frac{N}{2}$, and $N = 2^p$.

The function, *s(t)*, is uniformly sampled at a frequency, f_s. The FFT frequencies, f_ℓ, are fixed by the sampling frequency, f_s, and the number of samples, *N*; in particular, $f_0=0$, $f_1=f_s/N$, plus $\frac{N}{2}-1$ harmonics. Importantly, this choice of f_ℓ (with $N = 2^p$) ensures an integral number of periods of the FFT mode within the sampled data.

Being constrained to these frequencies allows efficient calculation of coefficients for large *N*. The FFT is efficient because it takes only $O(N \log_2 N)$ operations to calculate these Fourier coefficients at $\frac{N}{2}+1$ frequencies. In particular, repeated evaluations of $e^{2\pi i \ell f_s/N t_k}\, s(t_k)$ occur for different frequencies, but the algorithm elegantly orders operations to eliminate these repetitions. Further, evaluation of the trigonometric functions in $exp(\, 2\pi\, i\, l\, f_s/N\, t_k\,)$ is simplified.

C. FFT Leakage and the Burgess-Lanczos Method

Yet another way to view leakage is to calculate the Fourier coefficients, C_ℓ, of a single mode, $\cos 2\pi f_\mu t$, where $f_\mu = \mu\, f_s / N$. Using (3), leakage in the continuous Fourier transform can be expressed in closed form as,

$$C_\ell = \frac{2}{T}\int_0^T \cos 2\pi f_\mu t\, \cos 2\pi f_\ell t\, dt \;\; ; \mu \neq \ell \qquad (5)$$

$$= \frac{\sin 2\pi(\mu+\ell)}{2\pi(\mu+\ell)} + \frac{\sin 2\pi(\mu-\ell)}{2\pi(\mu-\ell)} \approx \frac{\sin 2\pi(\mu-\ell)}{2\pi(\mu-\ell)}\,.$$

For integral μ, the mode $\cos 2\pi f_\mu t$ is periodic with period *T*, and $C_\ell = \delta_{\mu\ell}$, that is, there is no leakage (δ is the Kronecker delta). For real μ between ℓ and $\ell+1$, that is, $\mu = \ell + \varepsilon$, for a frequency perturbation ε, the mode is not periodic with period *T*. Non-zero Fourier coefficients occur across the spectrum. Fig. 3 graphs (5) for the modes in Fig. 1 and compares with FFT results.

The Burgess-Lanczos method circumvents leakage by estimating a mode from adjacent FFT modes. It assumes the signal contains a single mode, $\cos 2\pi f_\mu t$, and uses (5) to approximate the unknown mode's frequency, magnitude, and phase. In particular, the adjacent Fourier (or FFT) coefficients $\hat{C}_\ell$ and $\hat{C}_{\ell+1}$ are assumed to dominate ($\ell \le \mu < \ell+1$), as the other coefficients quickly become small. From FFT results, the mode may be approximated [1] by,

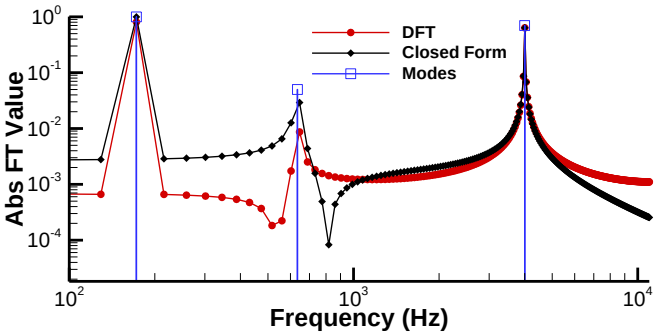

Fig. 3. Graphs of closed-form transform (5) and an FFT (*N*=512) for the three modes of the signal in Fig. 1, by superposition.

$$\varepsilon = \frac{\left|\hat{C}_{\ell+1}\right|}{\left|\hat{C}_\ell\right| + \left|\hat{C}_{\ell+1}\right|}\;; \qquad f_\mu = (\ell+\varepsilon)\frac{f_s}{N}\;;$$

$$\left|\hat{C}_\mu\right| = \frac{2\sin(\pi\varepsilon/N)}{A_r \sin\pi\varepsilon}\left|\hat{C}_\ell\right| \qquad \theta_\mu = \theta_\ell + \frac{\pi(N-1)\varepsilon}{N} \qquad (6)$$

where A_r is a normalization, typically $1/N$, sometimes 1. Equations 6 are appropriate for $0 \le \varepsilon \le 0.5$, but, for $0.5 < \varepsilon < 1$, better estimates of C_μ and θ_μ are:

$$\left|\hat{C}_\mu\right| = \frac{2\sin(\pi(1-\varepsilon)/N)}{A_r \sin\pi(1-\varepsilon)}\left|\hat{C}_{\ell+1}\right|;$$

$$\theta_\mu = \theta_{\ell+1} - \frac{\pi(N-1)(1-\varepsilon)}{N}\,. \qquad (7)$$

Burgess [1] also indicated that Hanning filtering and larger sample sizes, *N*, would give improved estimates, and he provided equations and error estimates. This method is very efficient and significantly improves the accuracy of FFT mode estimates for low noise signals. Prediction accuracy is degraded by the leakage of nearby modes and by noise. Fig. 3 suggests accuracy limitations, since the closed form and FFT results differ, as do FFTs of different lengths, *N*; further these differences are comparable in size to the small mode.

II. ALGORITHM OUTLINE

To accurately determine Fourier coefficients for a short duration signal, the current algorithm uses NLLS and a classical search strategy, namely bracketing and search, to minimize a function and find a mode. A third step, polishing, is also required. The function to be minimized, $\Phi(S_k)$ (8), is the power of a reduced signal, namely, the signal, S_k, minus a cancelling mode. The search space includes frequency, f_{test}, magnitude, C_{test}, and phase, θ_{test}.

$$\Phi(S_k) = \frac{1}{N}\sum_{k=0}^{N-1}\left[S_k - C_{test}\cos(2\pi f_{test} k\Delta t + \theta_{test})\right]^2\,. \qquad (8)$$

Bracketing must exclude local minima from the search.

Powell's method is the most efficient and robust search method tested. Once $\Phi(S_k)$ is minimized, the cancelling mode is subtracted from the signal, S_k, to form a new residual signal, S_k^M, (9), and the bracketing and search is repeated with $\Phi(S_k^M)$ to find the next largest mode. These steps are shown in Table I.

$$S_k^M = S_k - \sum_{j=1}^{M} C_j \cos(2\pi f_j k\Delta t + \theta_j) \quad k = 0, N-1 \tag{9}$$

Just as the Burgess-Lanczos estimate has errors due to interaction with other modes, the search result is not exactly the mode of the signal. In particular, the minimum of $\Phi(S_k)$ is shifted away from the exact result by the presence of other modes in the signal, hence the frequency, magnitude, and phase found for the cancelling mode are well approximated, but not exact. To circumvent this problem, as many of these offending modes as possible are removed while leaving the mode to be refined. In particular, the polishing function S_{kJ} (10) isolates mode J by subtracting all other approximated modes from the original signal leaving only the residual signal and the Jth mode to be polished. The search method is also applied to this isolated mode, $\Phi(S_{kJ})$.

$$S_{kJ} = S_k - \sum_{j \neq J} C_j \cos(2\pi f_j k\Delta t + \theta_j)$$

$$\text{for } k = 0, N-1, \quad j = 1, M, \ j \neq J. \tag{10}$$

The algorithm continues until the signal is sufficiently well represented. The algorithm is implemented in single precision, P=32. Table II demonstrates results and compares with the Burgess-Lanczos method.

III. Analysis of Bracketing, Search Methods, and Polishing

This section provides more detailed explanations and analysis of bracketing, search, and polishing. Two important properties must be guaranteed. First, ensure generally that bracketing includes only the region of attraction of a global minimum—no local minima. Second, ensure that polishing works generally, and the global minimum converges to the desired mode. For a first reading, scrutiny of all details may not be necessary.

TABLE I
OUTLINE OF ALGORITHM AND PERFORMANCE

	% CPU Time	Operation Count
Form discrete signal, $S_k = s(k/f_s)$, $k=0,N-1$		
For each cancelling mode j, $j = 1, M$		
{		
FFT of signal, S_k^{j-1}, for Initial Bracketing	*2.1*	*O(M N log₂N)*
For each test frequency, f_l, , $l = 1, B_f$		
{ DFT_F (f_l) } /* final bracketing */	*4.5*	*M B_f 10 N*
Refinement search by Powell's method $\Phi(S_k^{j-1})$	*28.5*	*O(M·10N)*
Remove Approx Mode, Reduced Signal, S_k^j	*< 0.1*	*M 8N*
/* Polishing */		
For each mode found, mode J, $J = 1, j$ {		
Form Polishing Signal, S_{kJ}, for mode J	*<0.1*	*O(M^2 N)*
Refinement search by Powell's Method $\Phi(S_{kJ})$	*64.8*	*O(M^2 10N)*
}		
}		

% CPU Time is based on signal in Fig 1; $N=512$, $M=3$, $B_f=21$.

A. Analysis of Bracketing

With an oscillatory signal and cancelling modes, many local minima are observed for Φ. Bracketing must exclude local minima and find a starting point within the region of attraction of the global minimum. However, perturbation analysis shows that an FFT helps find the bracketing interval.

1) Perturbation Analysis of the Global Minimum

Perturbation analysis of the bracketing requirements considers a simplified model problem, namely, a single mode, $\cos(2\pi ft)$, representing the discrete target signal, $S_k = \cos(2\pi k f/f_s)$. Further, the analyses involve an approximate cancelling mode with a small error in frequency, magnitude, or phase. Table III shows these cancelling modes and error parameters. The minimization function is the substitution of the discrete target signal and cancelling mode into Φ, (8). Note that the summation in Φ is approximated by the integral $f_s \int_0^T dt$.

Fig. 4 and Fig. 5 show these perturbation expansions in frequency, magnitude, and phase neighborhoods of the global minimum. A spreadsheet calculation, with the signal and Φ (8) was used to validate the perturbation analysis, and it is also plotted in the figures.

TABLE II
Performance Comparison

	f_1	C_{f1}	θ_{f1}/π	f_2	C_{f2}	θ_{f2}/π	f_3	C_{f3}	θ_{f3}/π
Exact	172.2656	1.0	0.3	4000	0.7	0.5	635	0.05	0.2
Current	172.2659	1.00000	0.30000	3999.9902	0.69999	0.500152	634.9979	0.05	0.20004
BL	172.3051	0.99918	0.29909	4000.0117	0.70026	0.49910	635.2963	0.050151	0.19431
BL Hanning	172.3492	0.99806	0.29807	3993.411	0.5468	0.6532	634.9552	0.049926	0.20153

Performance comparison of the current search algorithm and the Burgess-Lanczos (BL) method with and without Hanning filtering on the noiseless, three-mode signal from Fig. 1. Burgess-Lanczos method modes successively removed, largest to smallest, to reduce interaction error; polishing was not advantageous. P = 32 bits.

TABLE III
PERTURBATION ANALYSIS OF REGION OF ATTRACTION

	Cancelling Mode	Perturbation Parameter	Validity	Dominant Term in Φ Expansion
Frequency	$cos(2\pi f(1+\varepsilon)t)$	$\varepsilon = \Delta f/f$	$\lvert\varepsilon\rvert < 1/(2\pi T)$	$\varepsilon^2(n^2/6)$; $\Phi(\Delta f) \sim \Delta f^2 T^2$
Magnitude	$(1+\delta)cos(2\pi ft)$	δ	exact	$\Phi(\delta) \sim \delta^2$ (1/2)
Phase	$cos(2\pi ft+\phi)$	ϕ	$\lvert\phi\rvert < 0.5$ rad, (30^o)	$\Phi(\phi) \sim \phi^2$ (1/2)

Perturbation analysis shows that the search for signal $cos(2\pi ft)$ has a quadratic global minimum for errors in frequency, ε, magnitude, δ, and phase, ϕ. Further, the bracketing interval for phase and magnitude is independent of parameters T and N. The bracketing interval for frequency depends on sampling interval period, T. $n=2\pi fT$, the number of mode periods in the sampling interval, times 2π.

As a guide to bracketing the search, Φ is quadratic in a neighborhood of the global minimum for errors in frequency, ε, magnitude, δ, and phase, ϕ. Further, the region of attraction for phase and magnitude is independent of sampling interval parameters T and N (Fig. 5).

The bracketing interval width for frequency depends on sampling interval period, T, and Fig. 4 shows two different periods, T=23.22 ms and T=11.61 ms. In particular, replacing ε in Table III with frequency error, Δf, the dominant quadratic term is $\Phi(\Delta f) \propto \Delta f^2 T^2$. Hence for shorter periods, T, Φ has a wider parabola near the minimum, and the frequency interval of attraction is larger.

This interval of attraction analysis is consistent with Angeby's observation [16] and implementation where short sliding windows significantly improve numerical properties.

2) Initial Bracketing by FFT Periodogram

The discrete modes of an FFT are separated by frequency intervals of f_s/N. The first bracketing step is to use an FFT periodogram to narrow the search. In particular, the largest FFT coefficient, $|C_f|$, is found, where $f = k\ f_s\ /N$, and a conservative frequency bracketing interval is $[(k-1)f_s/N, (k+1)f_s/N]$—a reduction to $4/N$ of the FFT's spectrum. A second bracketing step provides further reduction to a starting point well within the global minimum's region of attraction.

3) Final Bracketing by Individual Mode DFT_F

The Burgess-Lanczos method could be used for final bracketing, but it is used for validation instead. An alternate method is introduced here, designated DFT_F.

The FFT is constrained to find Fourier coefficients at frequencies separated by f_s/N, and this separation is determined by the sampling frequency, f_s, sample size, N, and the requirement for an integral number of periods of the mode. Yet, both testing between the FFT frequencies, f_{test}, and successful bracketing is possible by using less that the full data set S_k, $k=0, N-1$, (Fig. 6) and by using an approximation to the Fourier Transform integral (4).

For a signal, $s(t)$, where the FFT suggests a spectral maximum near $k\frac{f_s}{N}$, consider a test mode at the frequency $f_{test} \in \left[(k-1)\frac{f_s}{N}, (k+1)\frac{f_s}{N}\right]$. The largest number of complete periods of this mode within the available N samples is $\lfloor f_{test} T \rfloor$, the number of data points inside this interval is $\tilde{N} = \frac{\lfloor f_{test} T \rfloor}{f_{test} T} N$

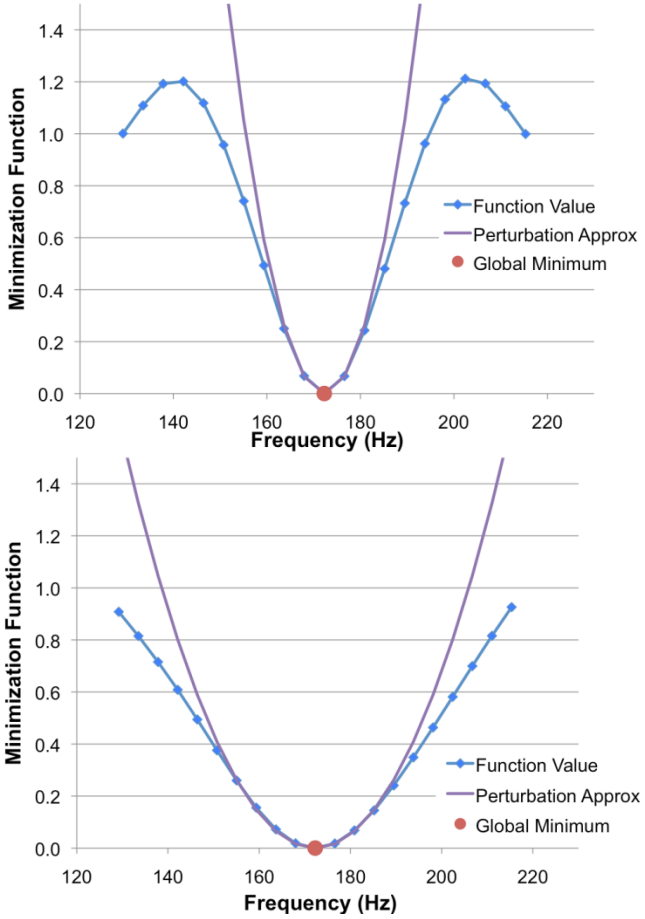

Fig. 4. Value of Φ and the perturbation expansion in a frequency neighborhood, $[(k-1)f_s/N, (k+1)f_s/N]$, for a single mode signal. The minimum is quadratic, and, importantly, the bracketing interval depends on the sampling period, T; upper, T=23.22 ms (N = 512), lower, T=11.61 ms (N= 256). The mode is f = 172.2656 Hz, C_f = 1.0, $\theta_f = 0.3\pi$. Magnitude, δ, and phase, ϕ, errors are zero.

, and the interval length is approximately $\tilde{T} = \frac{\lfloor f_{test} T \rfloor}{f_{test} T} T$ in

$$\hat{C}_{f_{test}} \approx \frac{2}{\tilde{T}} \sum_{k=0}^{\tilde{N}-1} e^{2\pi i f_{test} t_k}\, s(t_k), \text{ where } t_k = \frac{k}{f_s}. \qquad (11)$$

Since less than a test mode's period is lost from each interval, $[0, T]$, high frequencies, f_{test}, have much less sample loss than low frequencies. Complete loss occurs and $\tilde{T}$ is zero for $f_{test} < f_s/N$; a brute force method is used for bracketing at these low frequencies.

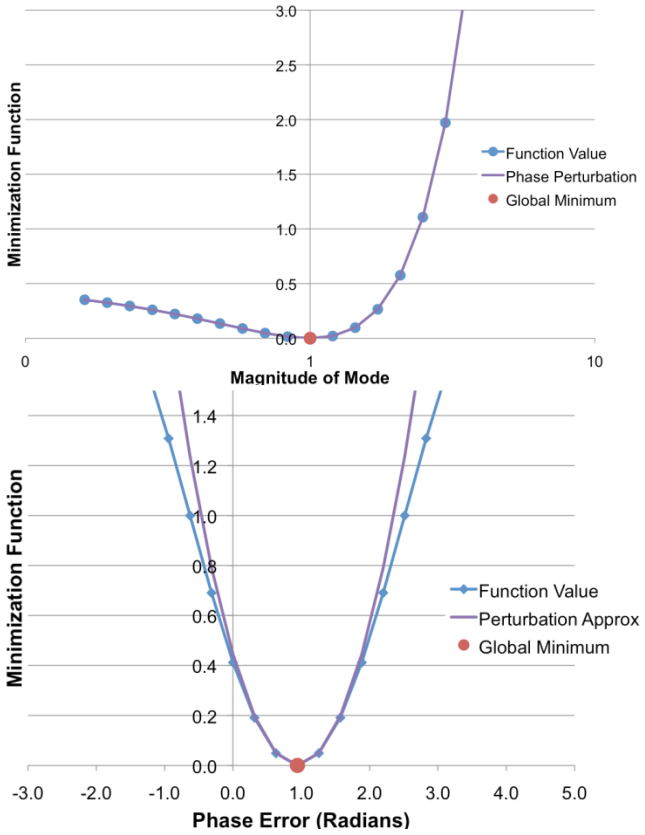

Fig. 5. Value of Φ and the perturbation expansion in magnitude (upper) and phase (lower) neighborhoods for a single mode signal. Perturbation analysis indicates each global minimum is quadratic and, importantly, the interval of attraction width does not depend on sample size, *N*.

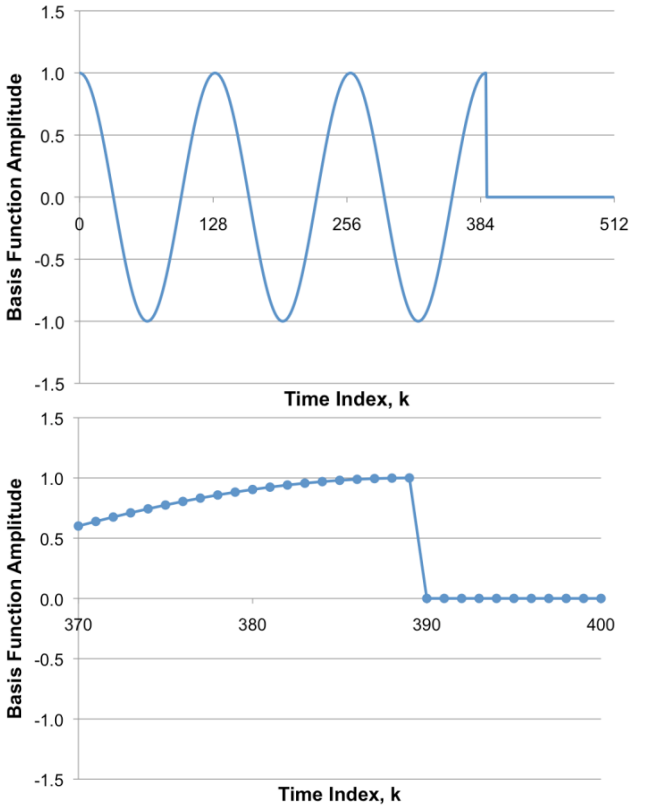

Fig. 6. Mode f_{test}=170 Hz for DFT_F on an interval, N=512. Less than all data points are used to ensure an integral number of periods of the mode.

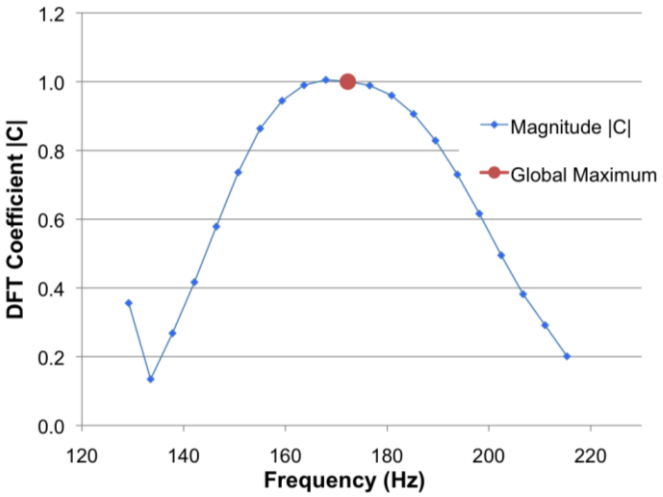

Fig. 7. Plot of DFT_F coefficient, $|C_{ftest}|$, in the interval [$3\ f_s/N$, $5\ f_s/N$].

The computational cost is estimated as $10N$ for each mode tested, and DFT_F relative cost is shown in Table I. The DFT_F shows leakage, but does not show minimum shifting.

To bracket the global minimum of Φ, the strategy is to evaluate the DFT_F at test frequencies, f_{test}, in the interval $[(k-1)f_s/N,\ (k+1)f_s/N]$. The DFT_F gives magnitude, C_{ftest}, and phase, θ_f. Fig. 7 shows an example. The largest DFT_F coefficient, $|C_{ftest}|$, approximates the center of the interval of attraction, and is used as the starting point for multidimensional minimization.

How close is close enough? Could the presence of other signal modes reduce the interval of attraction from the single mode estimates of Table III? Experimentally, no issues have been detected in a large number of signals after searching to a resolution of 0.1 f_s /N. Theoretically, the multimode perturbation analysis of section C shows that other modes do not change the quadratic expansion term—only the minimum's location is shifted.

B. *Multidimensional Minimization by Powell's Method*

The search algorithm must efficiently and exactly perform multidimensional minimization on $\Phi(S_k)$, that is, find the frequency, f, magnitude, C_f, and phase, θ_f, of a mode that best cancels a mode in the signal, S_k. Other search algorithms have been used to verify the code, yet Powell's method is robust, can provide quadratic convergence (the current implementation [11] does not include the necessary extensions), and requires only function evaluations—not derivative evaluations.

C. *Perturbation Analysis of Polishing Requirements*

Perturbation analysis shows that while searching for a cancelling mode by minimizing $\Phi(S_k)$, the minimum is shifted away from the correct cancelling mode by the presence of other modes in the signal. Hence the frequency, magnitude, and phase of the cancelling mode are well approximated by searching for the minimum of Φ, but polishing is required for exact results.

TABLE IV
MULTIPLE MODES REQUIRE POLISHING

	Cancelling Mode	Error Parameter	Validity	Error in Minimum
Frequency	$cos(2\pi f(1+\varepsilon)t)$	$\varepsilon = \Delta f/f$	$\lvert\varepsilon\rvert < 1/(2\pi T)$	$3a/n^2\ E_1$
Magnitude	$(1+\delta)\ cos(2\pi ft)$	δ	exact	$a\ E_2$
Phase	$cos(2\pi ft+\phi)$	ϕ	$\lvert\phi\rvert < 0.5$ rad, (30^o)	$a/2\ E_3$

Perturbation analysis shows the search minimum is shifted away from the desired mode by other modes in the signal. $\alpha = 2\pi f$, $\beta = 2\pi f_2$. Equation (12) gives the terms E_1, E_2, and E_3. $n=2\pi fT$, the number of mode periods in the sampling interval, times 2π, hence errors are small.

The extended perturbation analysis uses two modes, $\cos(2\pi ft) + a\cos(2\pi f_2 t)$, to represent the discrete target signal, S_k. The parameterization of the cancelling mode is as before.

Unfortunately, Φ is not of the form $A\varepsilon^2$ near the minimum as for a single mode. Instead, Φ is of the form $A\varepsilon^2 + B\varepsilon + C$ in a neighborhood of its minimum at $\varepsilon = -B/(2A)$ where A, B, and C are expressions that are constant in a neighborhood of the minimum. In particular, the presence of a second mode shifts the minimum of Φ away from $\varepsilon = 0$ and the desired result. Similarly for δ and ϕ.

These errors are small. Table IV and (12) show the error in the minimum when the signal and cancelling mode are substituted into Φ. For the signal in Fig. 1 without polishing, the frequencies are correct to 0.16 Hz, the magnitudes to 0.4%, and the phases to within 0.2 degrees.

$$E_1 = \frac{\alpha}{\alpha+\beta}\cos(\alpha+\beta)T + \frac{\alpha}{\alpha-\beta}\cos(\alpha-\beta)T - \frac{\alpha}{\alpha+\beta}\frac{\sin(\alpha+\beta)T}{T(\alpha+\beta)} - \frac{\alpha}{\alpha-\beta}\frac{\sin(\alpha-\beta)T}{T(\alpha-\beta)}$$
$$E_2 = \frac{\sin(\alpha-\beta)T}{T(\alpha-\beta)} + \frac{\sin(\alpha+\beta)T}{T(\alpha+\beta)} \qquad (12)$$
$$E_3 = \frac{\cos((\alpha-\beta)T)-1}{T(\alpha-\beta)} + \frac{\cos((\alpha+\beta)T)-1}{T(\alpha+\beta)}$$

To circumvent this problem, a discrete polishing signal, S_{kJ}, (10) is formed where all approximated cancelling modes are removed—except for the focus of the polishing, the J^{th} cancelling mode. By removing the best estimates of the other modes in S_{kJ}, this interference is reduced. By repeated polishing, the coefficients, C_f, of the cancelling modes converge to the best representation of the discrete signal.

IV. Important Algorithm Properties

Several algorithm properties are important and are compared in this section, namely, convergence, computational complexity, convergence metrics, tolerance of signal noise, and resolution of two modes with nearly identical frequencies.

A. Convergence to the Signal

Fig. 8 shows convergence versus computation time, that is, the reduction in error with additional Powell search iterations. The graph indicates linear convergence. Powell's method *can* give quadratic convergence, but the implementation used here [11], does not include the necessary extensions.

B. Computational Complexity

The current algorithm must be computationally efficient for all plausible input. Table I gives the observed relative CPU time for the signal in Fig. 1, plus the operation count for each algorithm step. The Powell's method searches and repeated evaluations of Φ are the most time consuming steps. The current algorithm is much more time consuming than a single FFT—50 to 100 times more in this example.

However, comparisons can be difficult. When performing an FFT, the temptation exists to increase the FFT's sample size, N, to increase the frequency resolution. Further, an FFT may need to be averaged over many sub-intervals [13] which further increases the computational cost. In contrast, the current algorithm is unencumbered by fixed FFT sampling frequencies and is accurate on a single interval; the temptation is to reduce sample size, N, and reduce computational cost. If

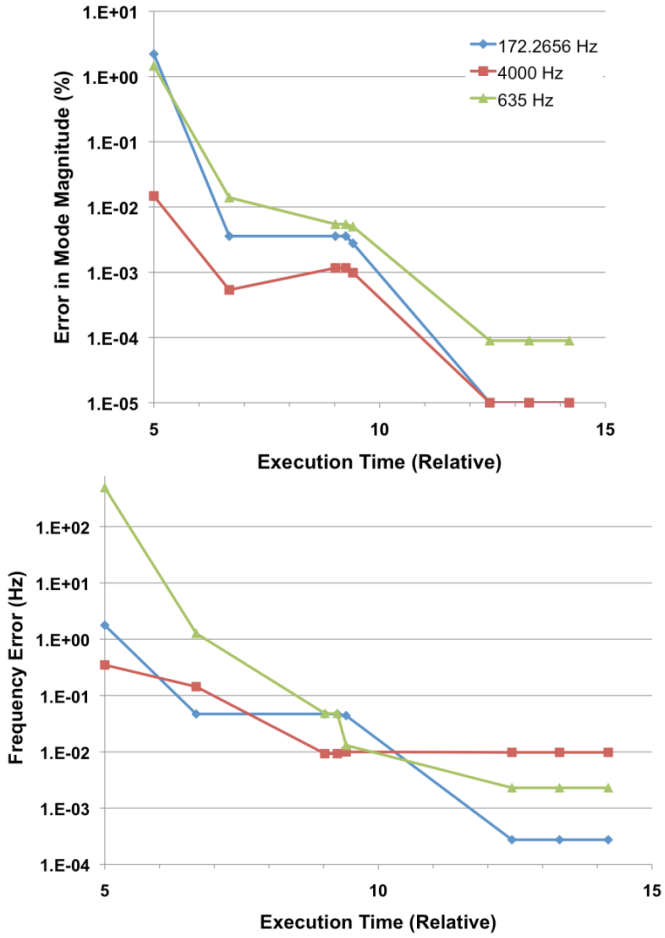

Fig. 8. Convergence of search versus computation time for the three mode signal of Fig. 1. Linear convergence is indicated. Relative time is the time to process a large number of identical signals. Beyond t=12, machine precision has been reached.

the Burgess-Lanczos method can successfully find signal modes, its computational efficiency is exceptionally good.

If noise or a fricative is present, if nearly identical frequencies are present, then the signal is not well captured by a few modes. Additional computations are required to find these additional modes. Termination conditions can help.

C. Precision and Convergence Metrics

The precision, P, of signal sample values, S_k, places limits on the smallest modes that should be searched for. In particular, audio formats often decode signal values to a precision $P = 8$ bits, and for a signal scaled so as not to exceed a magnitude of one, signal details smaller than 2^{1-P} are lost by rounding to the nearest bit. The search algorithm should not search for modes smaller than this precision threshold. Similarly, signal peak magnitudes exceeding 1.0 are rounded to one. These two limits define the dynamic range of a signal.

Using this concept, two convergence metrics are readily calculated to terminate the mode search. A convergence metric, L_1, compares the extreme value of the residual signal, S_k^M, to the precision threshold. Values less than 10 dB are considered good.

$$L_1 = 20\log(\underset{k=0,N-1}{Max}\left|S_k^M\right| / 2^{1-P}) \tag{13}$$

A second metric, L_2, compares the residual signal's power to the precision threshold.

$$L_2 = 20\log(\sqrt{\frac{1}{N}\sum_{k=0}^{N-1}(S_k^M)^2} / 2^{1-P}) \tag{14}$$

A third metric, L_3, measures the reduction in signal power due to removing cancelling modes.

$$L_3 = 10\log(\sum_{k=0}^{N-1}(S_k^M)^2 / \sum_{k=0}^{N-1}(S_k^0)^2) \tag{15}$$

Note, signals that do not utilize the dynamic range of the precision will have poor L_3 performance. Further note that for each metric, larger negative values indicate better performance. Table V shows metric values for each test case.

For storage, audio formats may use signal compression to increase signal precision, P, and/or store in fewer bits. For example, current telephone signal compression codes 12 or 13 bits into an 8-bit value, and current MP3 and Ogg Vorbis formats store in fewer than P=8 bits. CD quality sound is P=16 bits.

D. Effect of Signal Noise on Bisection, Search, and Results

Noise can introduce random errors into the predictions of mode frequency, magnitude, and phase. Further, noise places limits on the smallest modes that should be searched for.

Fig. 9 shows the three-mode signal, from Fig. 1, with and without Gaussian random noise, $r(t)$. In both cases, the three modes are found, however the predicted parameters have small errors. Further, these errors are random since mode searches in different sampling intervals will have different noise signals. Lastly, the noise is large enough—relative to the convergence criteria—that the algorithm adds cancelling modes in an effort to capture this noise. The convergence criteria must consider the noise level to avoid this undesirable effect.

Fig. 10 shows the prediction errors at various noise levels for the current and Burgess-Lanczos methods.

E. Resolution of Two Modes with Nearly Identical Frequencies

Can the algorithm resolve modes with nearly the same frequency? Any spectral resolution algorithm must deal with the fundamental problem that: two modes at nearly the same frequency on a short interval look like a single mode at the mean frequency, if $O(2\pi\Delta f\, T) < 1$.

Consider two modes of frequency $f+\Delta f$ and $f-\Delta f$, separated by a small frequency spacing, $\lambda = \Delta f/f$, and a phase offset $2\phi \in [0, 2\pi]$. Then, for any metric/algorithm, the combined signal is:

$$\begin{aligned} &\sin(2\pi f(1+\lambda)t + \phi) + \sin(2\pi f(1-\lambda)t - \phi) \\ &\equiv 2\sin(2\pi ft)\cos(2\pi f\lambda t + \phi) \\ &\approx 2\sin(2\pi ft)\,[\cos(\phi) + \sin(\phi)\cdot 2\pi f\lambda t + \cos(\phi)\,2(\pi f\lambda t)^2 + \ldots]. \end{aligned} \tag{16}$$

For short duration signals, the sampling interval, $[0,T]$, is small, and these two modes approach a single mode when $O(2\pi\Delta f\, T) < 1$. Note that this general analysis applies to any algorithm designed to resolve frequency components. Since T

TABLE V
CONVERGENCE OF EXAMPLE SIGNALS

Test Case	Precision, P	Current Method			Burgess-Lanczos		
		L_1 (dB)	L_2 (dB)	L_3 (dB)	L_1 (dB)	L_2 (dB)	L_3 (dB)
No Noise, 3 Mode Case, Fig. 1	32	8.1	-5.0	-91.0	48.2	33.0	-53.9
Noise, 3 Mode Case, Fig. 9	32	59.1	45.8	-40.2	74.7	57.9	-28.1
Two Modes Δf=20Hz, Fig. 11	32	52.2	26.4	-60.5	Not Stable		
Two Modes Δf=50Hz, Fig. 12	32	22.2	0.9	-87.4	89.6	74.9	-13.4
Guitar Open e, Fig. 13	8	0.0	-14	-25.3	9.6	-3.0	-14.4
Voiced Phoneme 'oo', Fig. 14	8	-0.1	-11.9	-22.6	16.9	2.3	-3.2
Fricative 's', Fig.. 15	8	6.5	-5.2	-10.8	20.3	8.6	-2.5
Plosive 'p', Fig. 16	8	19.9	2.7	-19.8	32.3	22.6	-1.1
Plosive 'b', Fig. 17	8	20.7	2.1	-22.9	35.0	25.3	-0.3

Convergence of residual signal, S_k^M, in each test case for each metric. Both the current method and the Burgess-Lanczos method are compared. Larger negative metric values indicate better performance.

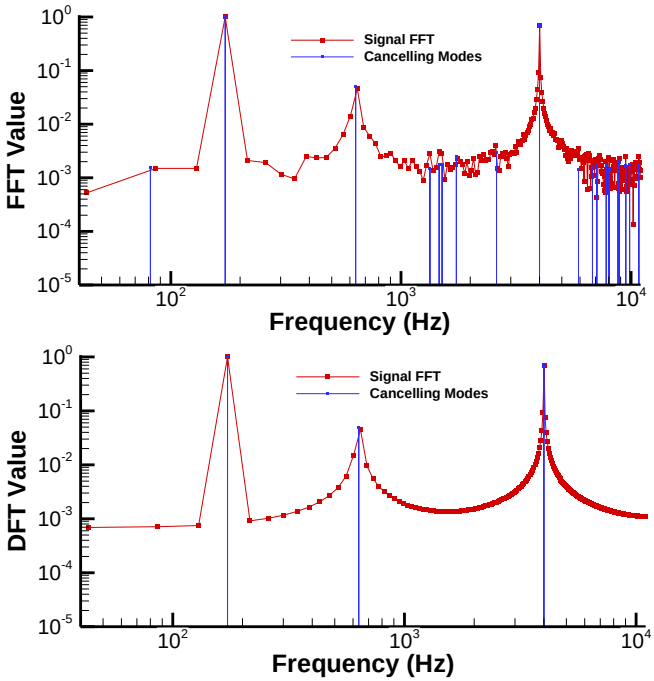

Fig. 9. Three mode signal (from Fig. 1) with Gaussian random noise (upper) and without (lower); SNR = 40 dB. The three modes are found, but additional cancelling modes occur. Leakage is still present for modes not at FFT frequencies.

$= N/f_s$, increasing the number of samples, N, or decreasing the sampling frequency, f_s, will improve resolution.

How does the current algorithm perform? Fig. 11 shows a three mode signal with two modes near 4500 Hz separated by 20 Hz, and $2\pi\Delta f\ T$=2.9. The two modes are not well resolved,

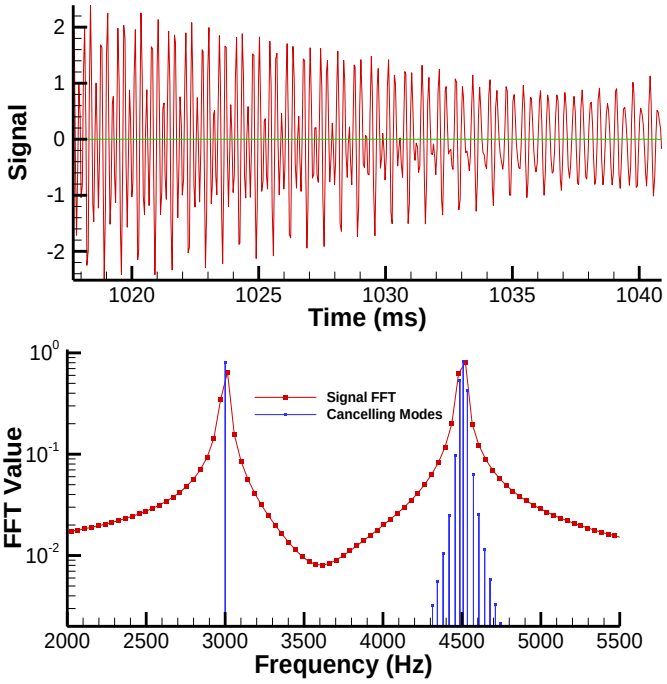

Fig. 11. Signal with two modes separated by Δf =20 Hz, and, lower, its FFT in red plus, the Fourier modes found, in blue. Signal (upper) consisting of three modes f = 4500 Hz, 4520 Hz, and 3000 Hz; C_f = 1.0, 0.9, 0.8. For $2\pi\Delta f\ T$=2.9, the two adjacent modes are not accurately resolved and additional spurious modes appear. The Fourier modes are a good approximation to the signal; residual signal shown in green (upper). Signal sampled at P =32 bits, M = 20.

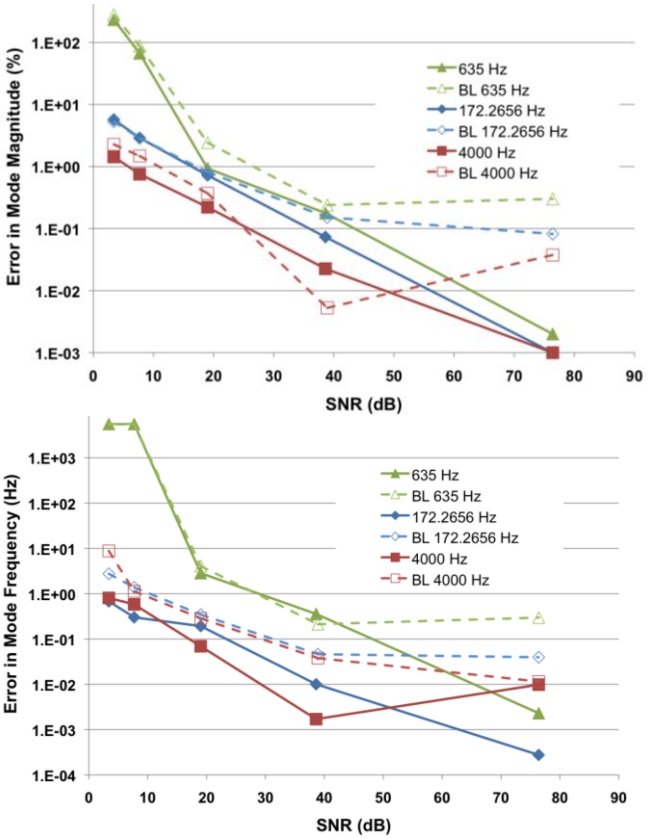

Fig. 10. Effect of noise level on precision performance of search algorithm and the Burgess-Lanczos method on the three-mode signal from Fig. 1.

and additional spurious modes are found. In contrast, Fig. 12 shows the identical signal except the two modes near 4500 Hz are separated by 50 Hz, and $2\pi\Delta f\ T$=7.3.

V. PRACTICAL EXAMPLES

Four auditory signals are presented in this section, and they range from a guitar note dominated by harmonics to voiced, fricative, and plosive phonemes from human speech. For each signal, results from the current algorithm are shown.

A. *Guitar Note*

The notes of string and woodwind instruments would be dominated by harmonics, like the guitar note in Fig. 13, due to the physics and mathematics of vibrating strings and air columns. Table VI shows modes found by the current method: fundamental frequency and eight harmonics plus two small, low frequency modes which are not harmonic. The residual signal, S_k^M, is below the precision bounds of the signal. Additional harmonics exist that are small and do not significantly reduce L_3. The Burgess-Lanczos method finds only the four largest modes, as shown in Table VI. The current method finds signal modes in this steady signal (Fig. 13) without spurious modes.

B. *Voiced Phoneme 'oo'*

A voiced phoneme 'oo' (vocal folds vibrating) from a male speaker is shown in Fig. 14. Table VII shows 5 harmonics of a fundamental frequency of 83.55 Hz, plus approximate

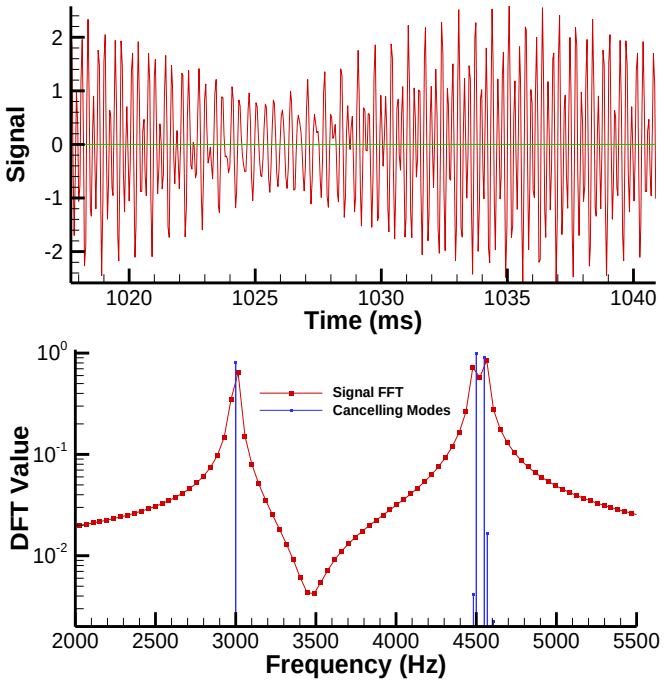

Fig. 12. Signal with two modes separated by Δf =50 Hz, and, lower, its FFT in red plus, the Fourier modes found, in blue. Signal (upper) consisting of three modes f = 4500 Hz, 4550 Hz, and 3000 Hz; C_f = 1.0, 0.9, 0.8. At $2\pi\Delta f\,T$=7.3, for the two adjacent modes, frequency, f, magnitude, C_f, and phase, θ_f, are found to about 1%; additional spurious modes are present, but they are small. Signal sampled at f_s = 22.05 kHz, N = 512, T = 23.22 ms, P =32 bits, M = 20.

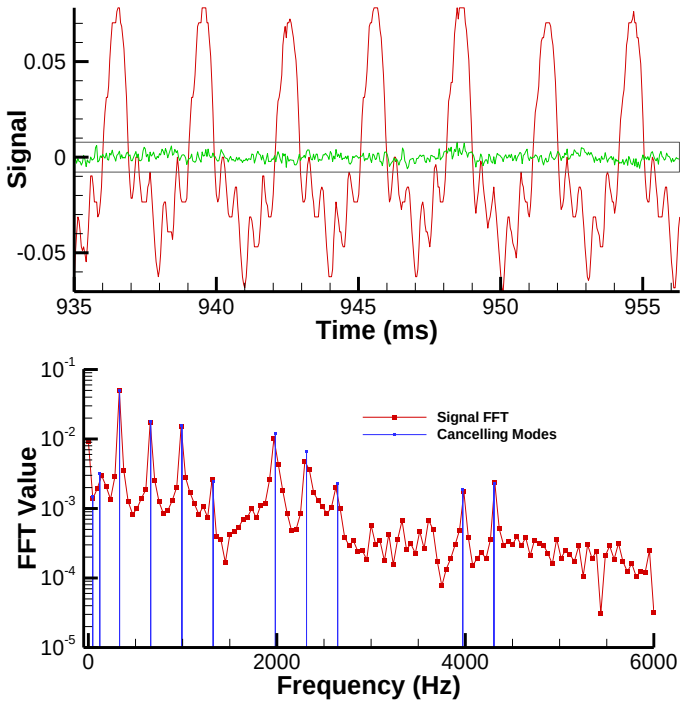

Fig. 13. Open e note of untuned guitar with fundamental frequency of 331.05 Hz and harmonics found using the current method. Upper graph is signal in red and residual signal, S_k^M , in green; the black horizontal lines indicate signal precision bounds. Lower graph is signal FFT in red, and cancelling modes as blue histogram lines. L_1=0.0 dB, L_2=-14 dB, L_3=-25.3 dB, P = 8. Properly tuned guitar open e is 329.63 Hz.

harmonics, found with the current method. The residual signal is below the precision bounds of the signal. The Burgess-Lanczos method finds only a single mode, as shown in Table VII.

The human vocal folds are a pair of semi-circular membranes restrained along their circular arcs; harmonics are known to occur. The shape of the human vocal tract is known to shift formant frequencies [24]; the approximate harmonics above 550 Hz in Table VII may be due to this effect. Modes are separated by more than 50 Hz except for three small modes.

C. *Fricative Phoneme 'S'*

A fricative phoneme, 'S', from a male speaker is shown in Fig. 15. Signal frequencies are centered near 4500 Hz. The weak signal is near the precision bounds of the signal. The modes found are typically separated by 50 Hz or more; only one separation of ~30 Hz exists. Increasing the number of cancelling modes, M, from 20 to 30 reduces L_1, L_2, and L_3 by only about 3 dB.

D. *Plosive Phoneme 'P'*

A plosive phoneme, 'P', from a male speaker is shown in Fig. 16. The stop (closed airway—no airflow or sound) is

TABLE VI
GUITAR OPEN E HARMONIC MODES

Current Method					Burgess-Lanczos		
Mode Freq (Hz)	**Harmonic**	**Magnitude**	**Relative Mag (%)**	**Phase (π Rad)**	**Mode Freq (Hz)**	**Magnitude**	**Phase (π Rad)**
46.00		0.0015	3.0	1.3814			
121.37		0.0032	6.4	1.6948			
331.05	1	0.0496	100.	1.4977	331.28	.0498	1.4916
661.43	1.998	0.0178	35.9	0.5794	662.86	.0178	.5472
992.10	2.997	0.0159	32.1	1.5536	992.02	.0157	1.5572
1323.95	3.999	0.0024	5.0	0.3348			
1983.22	5.991	0.0119	24.0	1.8695	1982.47	.0118	1.8887
2314.22	6.991	0.0067	13.5	1.9785			
2645.17	7.990	0.0023	4.6	1.0998			
3972.34	11.999	0.0019	3.8	1.7988			
4303.84	13.001	0.0023	4.6	1.3622			

Modes found by the two methods for the guitar open e note of Fig. 13.

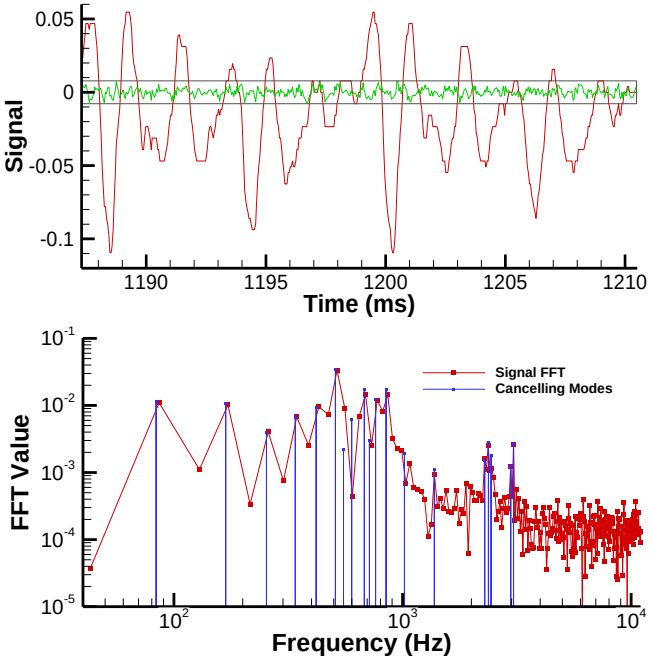

Fig. 14. Voiced phoneme 'oo' signal (upper) in red and residual signal in green; the horizontal lines indicate signal precision. In lower graph, signal FFT in red, cancelling modes as blue histogram lines. L_1=-0.1 dB, L_2=-11.9 dB, L_3=-22.6 dB.

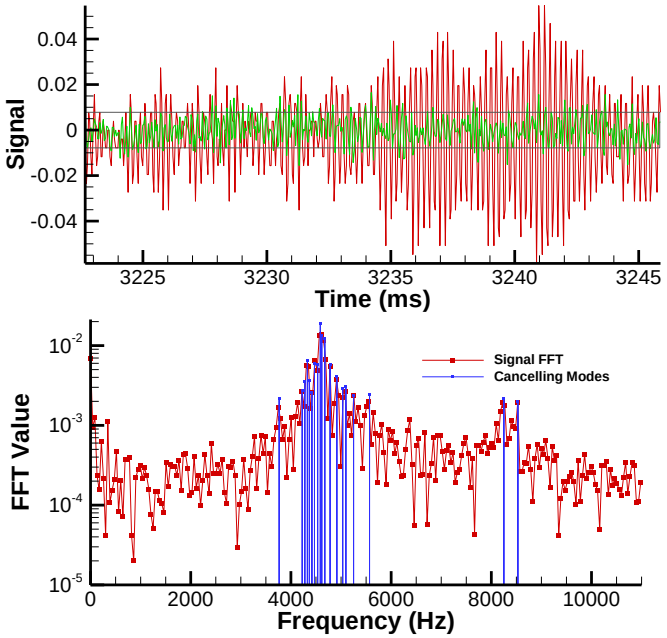

Fig. 15. Fricative phoneme 'S' signal (upper) in red and residual signal in green; the horizontal lines indicate signal precision. In lower graph, signal FFT in red, and cancelling modes as blue histogram lines. L_1=6.5 dB, L_2=-5.2 dB, L_3=-10.8 dB, M=20.

shown before 7605 ms. The strong pressure wave resulting from resumed airflow is clearly visible in the signal—and the frequency spectrum has large, low frequency modes. The fricative modes found near 2000 Hz are typically separated by about 50 Hz; three spacings are of ~30 Hz.

All signals were recorded in Ogg Vorbis format [25] using Audacity [26] on a PC with a Gigaware desktop microphone. The human speech samples are of a male speaker. All searches are limited to twenty modes, M=20. Signals are sampled at f_s = 22.05 kHz (recorded at 44.1 kHz and desampled), N = 512, T = 23.22 ms, and P = 8 bit resolution. Signals are assumed to be bandlimited even though the Nyquist frequency is 11 kHz and the frequency limit of human auditory sensitivity is 20 kHz.

VI. Conclusion

For the perception of tone in human speech and music, capturing the subtle variations in modal frequencies and magnitudes is important. The current method can accurately find the spectrum of these short duration signals.

Each of the algorithms considered here are useful. The FFT continues to be a pillar of signal processing. If better accuracy is required in a low-noise signal with a few modes, the Burgess-Lanczos algorithm is efficient and appropriate. Yet, if

TABLE VII
VOICED PHONEME 'OO' MODES

Current Method					Burgess-Lanczos		
Mode Freq (Hz)	**Harmonic**	**Magnitude**	**Relative Mag (%)**	**Phase (π Rad)**	**Mode Freq (Hz)**	**Magnitude**	**Phase (π Rad)**
83.55	1	.0114	33.3	0.3982			
168.38	2.015	.0106	31.0	0.7789			
253.88	3.038	.0040	1.2	0.8194			
339.30	4.061	.0069	20.1	0.7760			
418.73	5.012	.0094	27.5	0.8434			
508.28	6.084	.0342	100.	0.3580	525.82	.0360	1.9824
551.64	6.602	.0022	6.4	0.1217			
599.09	7.170	.0062	18.1	1.4409			
680.28	8.142	.0173	50.6	1.8398			
715.84	8.568	.0030	8.8	1.5003			
762.13	9.122	.0123	36.0	1.9616			
847.26	10.141	.0175	51.2	1.5075			
2373.52		.0028		1.3395			
3057.92		.0026		1.8191			

Modes found by the two methods for the voiced phoneme 'oo' of Fig. 14.

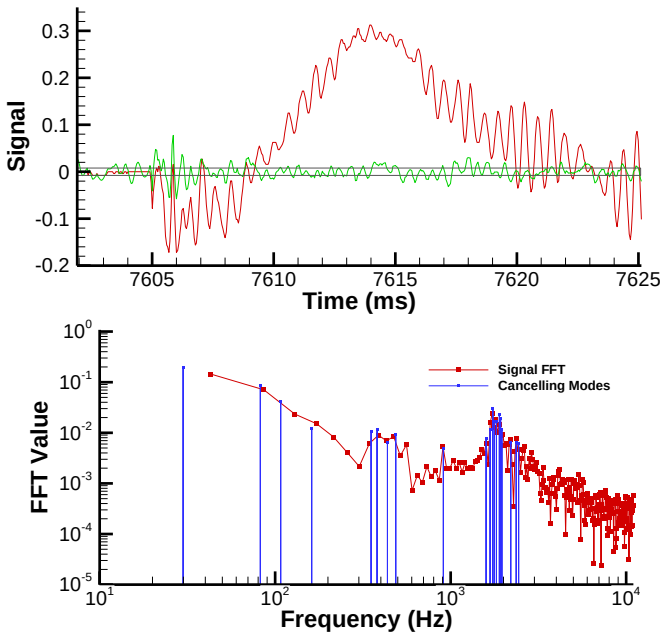

Fig. 16. Plosive phoneme 'P' signal (upper) in red, residual signal in green; the two black horizontal lines are the signal precision bounds, P = 8 bit resolution. In lower graph, signal FFT in red, and cancelling modes as blue histogram lines. L_1=19.9 dB, L_2=2.2 dB, L_3=-19.8 dB..

high accuracy Fourier coefficients are required in a short duration signal with many modes, the current algorithm would be more appropriate.

ACKNOWLEDGMENT

Discussions with Donald Braun provided valuable insights into the Burgess-Lanczos method. Ambady Suresh provided very helpful comments on the paper.

REFERENCES

[1] J. C. Burgess, "On Digital Spectrum Analysis of Periodic Signals," *J Acoust Soc Am*, vol. 58, no. 3, pp. 556-567, Sept 1975.

[2] H. L. F. Helmholtz, On the Sensations of Tone as a Physiological Basis for the Theory of Music, 6th English Edition (1st German Edition 1862) ed., New York: Peter Smith, 1948.

[3] J. Fourier, The Analytic Theory of Heat, Paris: Firmin Didot Pere et Fils, 1822.

[4] A. A. Michelson and S. W. Stratton, "A new Harmonic Analyzer," *American Journal of Science*, vol. 5, no. 25, pp. 1-13, Jan. 1898.

[5] W. Thompson, "The tide gauge, tidal harmonic analyser, and tide predicter," *Proceedings of the Institution of Civil Engineers*, vol. 65, pp. 3-21, 1881.

[6] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex Fourier series," *Math. Comput.*, vol. 19, no. 90, pp. 297-301, 1965.

[7] J. W. Cooley, P. Lewis and P. Welch, "Historical Notes on the Fast Fourier Transform," *Proceedings of the IEEE*, vol. 55, no. 10, October 1967.

[8] S. Welaratna, "Thirty Years of FFT Analyzers...," *Sound and Vibration*, January 1997.

[9] A. C. Gilbert, S. Guha, P. Indyk, S. Muthukrishnan and M. Strauss, "Near-Optimal Sparse Fourier Representations via Sampling," in *Proceedings of the 34th Annual ACM Symposium on Theory of Computing*, Montreal, Canada, 2002.

[10] A. V. Oppenheim and R. W. Schafer, "Power Spectrum Estimation," in *Digital Signal Processing*, Englewood Cliffs, NJ, Prentice-Hall, 1975, pp. 532-571.

[11] W. H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling, Numerical Recipes in C: The Art of Scientific Computing, New York: Cambridge Univ. Press, 1988.

[12] M. S. Bartlett, An introduction to stochastic processes: with special reference to methods and applications, New York: Cambridge University Press, 1953.

[13] D. G. Lewicki and J. J. Coy, "Vibration Characteristics of OH-58A Helicopter Main Rotor Transmission," NASA TP-2705, Cleveland, OH, 1987.

[14] C. Lanczos, Discourse on Fourier Series, New York: Hafner, 1966. p 139.

[15] B. Friedlander and J. M. Francos, "Estimation of Amplitude and Phase Parameters of Multicomponent Signals," *IEEE Trans. Signal Process.*, vol. 43, no. 4, pp. 917-926, Apr. 1995.

[16] J. Angeby, "Estimating Signal Parameters Using the Nonlinear Instantaneous Least Squares Approach," *IEEE Trans. Signal Process.*, vol. 48, no. 10, pp. 2721-2732, Oct. 2000.

[17] L. Rabiner and B.-H. Juang, Fundamentals of Speech Recognition, Upper Saddle River, NJ: Prentice-Hall, 1993.

[18] X. Huang, A. Acero and H. W. Hon, Spoken Language Processing: A Guide to Theory, Algorithm, and System Development, Upper Saddle River: Prentice-Hall, 2001.

[19] A. Grossmann and J. Morlet, "Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape," *SIAM J. Math. Anal.*, vol. 15, no. 4, pp. 723-736, Jul. 1984.

[20] S. G. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674-693, Jul. 1989.

[21] G. V. Békésy, "The variation in phase along the basilar membrane with sinusoidal vibrations.," *J. Acoust. Soc. Am*, vol. 19, no. 3, pp. 452-460, 1947.

[22] C. D. Geisler, From Sound to Synapse: Physiology of the Mammalian Ear, New York: Oxford Univ. Press, 1998.

[23] W. E. Boyce and R. C. DiPrima, Elementary Differential Equations and Boundary Value Problems, New York: John Wiley & Sons, 1977 p. 470-472.

[24] J. Sundberg, "The Acoustics of the Singing Voice," *Scientific American*, vol. 236, no. 3, pp. 82-91, Mar. 1977.

[25] Ogg Vorbis, "Vorbis.com: open source audio encoding and streaming technology," [Online]. Available: http://www.vorbis.com. [Accessed Feb. 24 2012].

[26] Audacity, "Audacity: Free Audio Editor and Recorder," [Online]. Available: http://audacity.sourceforge.net. [Accessed 1 January 2012].

Mark E. M. Stewart (M'2012) received the B. Math degree in applied mathematics from the University of Waterloo, Waterloo, Canada in 1985 and the M.A. and Ph.D. degrees in applied and computational mathematics from Princeton University, Princeton, NJ in 1987 and 1990.

Currently, he is a senior research engineer working at NASA Glenn Research Center in Cleveland, Ohio. His NASA work involves high-fidelity, scientific computing simulations of propulsion systems, including multidisciplinary fluid/thermal/structural analysis of nuclear thermal propulsion fuel elements, microgravity fluid flow, turbofan engines, hypersonic air-breathing engines, external aerodynamics, combustion, multidisciplinary fluid-combustion-thermal-structural analysis, and engineering design. His outside interests include semantic analysis of scientific programs and spreadsheets. The problem in the current paper arose in attempts to simulate neural processing of semantics in the human auditory system.

Dr. Stewart is a member of the IEEE and the AIAA.